ABSTRACT

Synchronization of multimedia data having at least audio and text sequences is disclosed. The audio sequence is divided into at least one audio data group, where a current audio data group is synchronized to a nearest time mark. The current audio data group is then associated to a number of a word in the text sequence corresponding to the current audio data group.